Pinning Alone?: A Study of the Role of Social Ties on Pinterest*

Changtao Zhong[1], **Nicolas Kourtellis**[2], **Nishanth Sastry**[1]

[1] Department of Informatics, King's College London, UK
[2] Telefonica I+D, Barcelona, Spain
[1] {changtao.zhong, nishanth.sastry}@kcl.ac.uk; [2] nicolas.kourtellis@telefonica.com

Abstract

This paper seeks to answer the question of whether social ties are important on interest-driven social networks, by analysing 4-years of activities of 50,000 randomly sampled users on Pinterest, a social image discovery website. We find that a non-trivial number of users' images are copied or repinned from strangers instead of friends, suggesting that social-based information exploration is not important. However, social interactions and social repins are critical for user retention: users interacting with friends are more likely to return Pinterest soon. These results suggest that the real role of social ties on interest-driven social networks is to enable bonding of users rather than seeking information.

1 Introduction

It has become *de rigueur* to create social networks amongst users on all kinds of Web 2.0 sites, especially those involving content sharing. Several prominent content- or product-driven sites, ranging from Pinterest (image-based sharing) and Vimeo (video sharing) to last.fm (music sharing) and Etsy (social shopping) incorporate social features such as the ability to follow other users activities, and to like or repost (share) content or products that they like. A key feature of social networks on content-driven sites is that links are intended to be made based on shared interests around an item or a category of items (Hendricks 2014; Jamison 2012).

However, recent research has shown that a majority of users do not participate in social aspects of content (Gelley and John 2015) or product (Swamynathan et al. 2008) sharing websites. Further, many users with explicit friendship or follow links may not in fact have any content that they like in common (Gelley and John 2015; Musial and Sastry 2012).

These findings stand in direct contrast to numerous studies where social networks have been shown to help in community formation, in a diverse range of interest- and goal-oriented environments and applications such as learning, working, medicine and online games. It has also been shown that borrowing links from existing mature social networks such as Facebook has helped content-driven social sites such as Pinterest and last.fm to *bootstrap* active and engaged communities (Zhong et al. 2014).

In light of these conflicting results, it is natural to ask what is the value of social networks on interest-based and content-driven sites. To answer this, we examine the complete activity history of ∼50k randomly selected users on Pinterest.com as a case study. We ask whether we can observe evidence for the different roles of social ties in interest-based social networks.

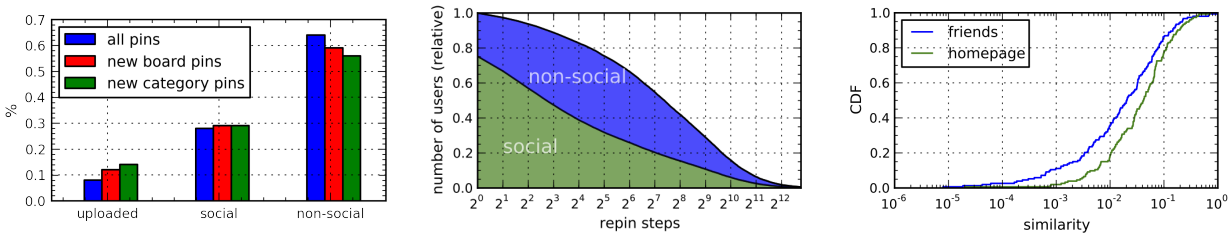
Findings. Surprisingly, for a goal-oriented and interest-based social network, a non-trivial proportion of information seeking happens through non-social means which indicates a decreased importance of social ties in content discovery on Pinterest. However, the social network still serves an important purpose for bonding: users who do engage with the social network, and in particular, users who have relatively close friends whom they know from offline contexts and from another social network (Facebook), are the most likely users to return to the platform.

In summary, these results show that, in fact, the previous literature results on interest-based social networks do not contradict each other. On Pinterest, as on other interest- and goal-oriented sites (Baird and Fisher 2005; Conole and Culver 2010; Heiberger and Harper 2008; DiMicco and others 2008; Lee and others ; Eysenbach 2008; Choi and Kim 2004; Ducheneaut and Moore 2004), the social network *does* play an important role in creating and fostering community formation. On the other hand, as claimed by others (Musial and Sastry 2012; Gelley and John 2015; Swamynathan et al. 2008), it is possible to 'get by' without it, and indeed, there appear to be more effective ways of finding information suiting one's interests, such as using the home page which features recent content.

2 Dataset and Terminology

Users on Pinterest.com collect so-called *pins*, which are images, together with associated URLs of webpages where they are found. The user who introduces an image into Pinterest is called its *pinner*; others who re-appropriate it for their own pinboards are *repinners*. Pins are 'pinned' onto a so-called 'pinboard' which is intended to be a thematic col-

*This work was partially supported by the Space for Sharing (S4S) project (Grant No. ES/M00354X/1) and by EU Horizon 2020 TYPES innovation action programme, No 653449.



(a) (b) (c)

Figure 1: **Social network is not critical for information seeking on Pinterest:** (a) *The source of pins in our dataset.* We examine whether the pins are from the users themselves ("uploaded"), repinned from friends of the user ("social"), or repinned from strangers ("non-social"). We term the first pins in users' boards as "new board pins", and the first pins of users in different categories as "new category pins". (b) *Long-term dynamics of social vs. non-social repins*, showing the proportion of social and non-social repins vs. the "maturity" of user on Pinterest, as measured by the number of repins made since joining. Note that there are fewer and fewer users as the "age" in terms of repin steps increases. (c) *Pins on homepage are more similar to a user's recent pins than to his friends' pins*. For this study we use repin activities from (Zhong, Karamshuk, and Sastry 2015). All of these pins were categorized using vectors constructed from 1000 objects detected on each pin, via Caffe (Jia et al. 2014). Cosine similarity was computed at random time points between vectors of pins of users and (1) pins of their friends (marked as "friends") and (2) pins of random users featured on the homepage at that time (marked as "homepage").

lection of related pins (e.g., one pinboard may have pins with images of different wedding dresses.). Most of pinboards are associated with one of 32 categories such as 'Design', 'Products', 'Home Decor', 'Animals and Pets', etc., which are globally recognised on Pinterest.

Pinterest incorporates social networking features to allow users to connect with other users with similar interests. Users can create connections to other users on Pinterest in two ways. The first is to explore the website and follow users they find interesting. We call social links created in this way *native links*, as they are created natively on the platform. The second way is using the "Find Friends" function. Users can connect their Facebook and Twitter accounts with their Pinterest accounts and the Friend Finder function will provide a list of Facebook and Twitter friends who are also registered on Pinterest. Users can select some of them to follow on Pinterest, and create *copied links*[1].

To understand the effects of the social network on pinning, we distinguish and quantify the effects of pins involving the social network of the pinner, which we call *social repins*, from *non-social repins*, which involve users to whom the pinner is not connected socially (i.e., repins of *non-friends* or *strangers*). We further distinguish between repins made from users who are friends with the original pinner as native and copied, based on the type of friendship link. Note that the nomenclature is always relative to the user who is (re)pinning an image onto their own pinboard.

Dataset. To analyse users' activities and understand how they accumulate and use social capital, we first created a dataset by crawling the entire pinning activity history of 50,000 users from the time they joined Pinterest until an arbitrarily chosen end date of April 1, 2014. To ensure that the sample was as unbiased as possible, the user IDs were randomly sampled from a near complete snapshot of the Pinterest social network collected and provided by Zhong *et al.* (Zhong et al. 2014) in Jan 2013. Thus, all users have at least 15 months of activity on Pinterest. Some of these user accounts have been suspended or deleted, and some have no pins. This left us with 48,185 users, who collectively have 3.9 million social links and 10.3 million pins.

3 Do Social Ties help Information Seeking?

A core function of interest-based and content-driven sites such as Pinterest is to enable users to find the information that suits their interests. Therefore, we might expect that the ability of social networks to provide access to new informationwould be important on Pinterest. To study this, we first compare social vs. non-social means of acquiring new information.

Previous studies (Zhong et al. 2013; Zhong, Karamshuk, and Sastry 2015) have found that creating new pins (or repins) is by far the most common activity on Pinterest, and presents a quintessential information seeking activity. Users may find out about the new pin externally to the website and upload the pin themselves. Or, they may repin an existing pin they find on Pinterest. In the latter case, they may repin a pin from a friend of theirs (i.e., a social repin), or they may repin an image pinned by someone with whom they are not connected socially (i.e., a non-social pin). We check what fraction of a user's pins come from each of the three sources – *uploaded pins, social repins and non-social repins*. In theory, each of these sources can be important to different extents in different kinds of information-seeking activities. For instance, at the time of the first pin of a user in a *new category*, the user may be less knowledgeable about that cate-

[1]We will focus on the connection between Pinterest and Facebook in this study, as more than 60% of Pinterest users have connected with Facebook accounts, while only about 10% of users connected to Twitter accounts (Zhong et al. 2014).

gory. Similarly, the first pin in a *new board* may be seen as a new 'sub category' or thematic collection, and may have different information seeking patterns, in comparison with subsequent pins.

We compare these different cases in Fig. 1a, across the 10.3 million images. Consistent with previous work (Zhong, Karamshuk, and Sastry 2015), we find that uploads are fewer in number than repins, whether social or non-social. However, the striking result is that across all kinds of information seeking, whether for the first pins in new boards/categories, or for subsequent pins, *non-social means of finding information from other users dominate over social repins.*

We then ask whether time matters: i.e., Do social repins become more important as the user matures and conducts more activities on Pinterest? Because the time between two pins may be widely different across users, we measure user age in terms of *repin steps*, the number of (re-)pins made since joining Pinterest. In Fig. 1b we examine the entire history of activities of all users, as they "age" in Pinterest by accumulating more activities. For each new repin activity of a user, we check whether it is a social or non-social repin. From the figure, we can see the proportion of social repins is much larger than non-social repins when users just join Pinterest. But the difference between the two proportions is reduced as users become more experienced in the platform. *This shows the growing importance of social repins for users with large numbers of repins, and for users just getting started on Pinterest (having very few repins), though a non-trivial portion of repin activities is still non-social.*

Such non-social repins are not hard to find: Pinterest highlights the most recent pins on the platform on its home page. Fig. 1c shows that on average, the recent pins being highlighted on the home page at a given point in time are likely to be more similar to user's current interests than the recent pins of the user's friends. Thus, given access to the homepage which proves to be a simple and easy to find source of interesting information, social-based information seeking becomes less critical.

4 The Real Social Benefits of Pinterest

Given the finding that social links are not critical for identifying pins, the most critical activity on Pinterest, it is puzzling that its social network is counted amongst the fastest growing across all platforms[2]. In this section, we address the issue of which users make social links, and why it is important. We find evidence the Pinterest social network is useful for bonding and interaction.

4.1 Social users are more likely to return soon

We start with a macro-scale analysis of users with different kinds of social links, and check whether the active users are also consistent, by measuring user retention. We follow Java *et al.* (Java et al. 2007), and mark users who have an activity in a given week as *active*. An active user is considered as *retained* if she also pins or repins in each of following X weeks. In our experiments, we measure the fraction of users retained among all active users in each week. Due to space constraints, we only show the results of $X = 3$, although similar results can be obtain for $X = \{1, 2, 4\}$.

[2]For example, see https://bitly.com/1MVk23k

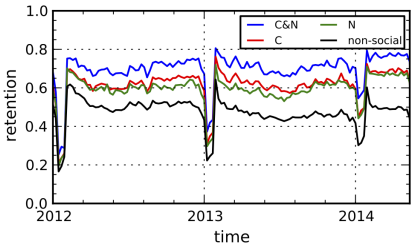


Figure 2: **Social users are more likely to return:** The retention rate of social users is significantly higher than for users who have no social repins. Among users with social repins, users who only have copied repins (C), i.e., those who have only repinned their Facebook friends, are more like to return than users who only have native repins (N), i.e., those who have only repinned friends made on Pinterest. Users with both types of repins (C&N) have the highest retention rate.

Fig. 2 shows the retention rates every week amongst different kinds of users. Apart from the annual drop in retention rates corresponding to the holiday season at the end of the year, retention rates for each kind of user stays at roughly the same level throughout the year. First, we divide active users in each week into social users and non-social users, according to whether they have made any social repins in the given week. These results show that retention of social users is higher than non-social users. We find that only about 50% of users who have not made any social repins ("non-social", black line) return to the platform within a week, whereas more than 60% of users that interacted with their social friends do return. This difference is statistically significant ($p = 3.15 \times 10^{-5}$). Thus we may conclude that users with social interactions are more likely to return and be engaged with the platform.

Furthermore, we divide social repins into copied and native repins according to whether friends are copied from Facebook or met natively in Pinterest as discussed in § 2. In Fig. 2, we can see that users who have a copied repin ("C", red line) have higher retention rate than users do native repins ("N", green line). Users with both types of repins ("C&N", blue line) show the highest retention rates. Users with both types of repins (21% of all users) contribute 71% of Pinterest activities.

4.2 Social sessions are less goal-oriented

It thus appears that social activities are important for the platform to retain users. Yet, social assistance appears to be less efficient than the homepage for individual users to find pins suiting their interests and goals. To understand the role of users' social interactions, we divide their pin timelines into "sessions", drawing session boundaries whenever there is a gap of T=6 hours or more between two consecutive pins. Thus, a session is all pins which are "close" to each other in time; similar results are obtained for other small values of T (e.g., 1 or 3 hours), but are not reported due to space.

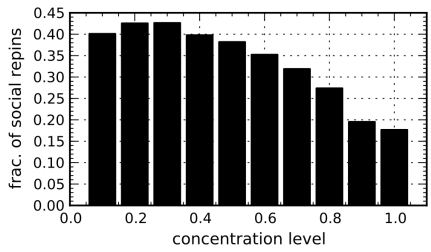


Figure 3: **Session concentration vs. social repin fraction:** Users tend to have more social repins in sessions without a goal (lower concentration) than in sessions with higher concentration.

In each session, we define the major category as the category into which the user has repinned the most number of images, and define the *concentration level* of this session as the fraction of images of the session that have been repinned to the major category. We expect that the higher the concentration level, the more likely users are to be searching for a specific category of information, and less likely to be browsing casually, without a goal. For each browsing session, we also compute the fraction of repins which are social. In Fig. 3, we compare the concentration level and social repin fraction. We notice that levels of less concentration are associated with higher levels of social repins, i.e., *when users do casual browsing without a specific information seeking need, they are more likely to be social, whereas when they have a specific information need, they are more goal oriented and find the information they need through the most efficient means (which may not be social)*. User interviews by Linder *et al.* (Linder, Snodgrass, and Kerne 2014) have also identified a similar pattern, with two types of information seeking behaviour in Pinterest: (a) casual browsing, when they do not have a particular goal in mind, and (b) specific searching, when users only respond to (repin) a specific type of image.

5 Discussion

In this paper, we have used the complete activity history of $\sim$50k randomly selected users to unpack the role and utility of the social network on Pinterest. We find that a non-trivial number of users' images are copied or repinned from strangers instead of friends, suggesting that social-based information exploration is not important. However, social users of Pinterest contribute the majority of activity, and have a higher probability of returning to the site. Social links have a strong bearing on engagement and this has been demonstrated in Facebook (Burke, Marlow, and Lento 2009), Twitter (Macskassy 2012), etc. The present work confirms this phenomenon on Pinterest as well, suggesting that bonding over social repins is important for the platform's functioning.

From the perspective of the platform operator and the collective community of users on the platform, the Pinterest social network is critical for healthy operation and to drive user activities. We may hypothesise that the social network could also be useful from an individual user's perspective, because of positive feelings of engagement and social interaction; however this requires further research (e.g., through user studies), to be confirmed.

References

Baird, D. E., and Fisher, M. 2005. Neomillennial user experience design strategies: Utilizing social networking media to support "always on" learning styles. *J. educ. tech. systems* 34:5–32.

Burke, M.; Marlow, C.; and Lento, T. 2009. Feed me: motivating newcomer contribution in social network sites. In *CHI*, 945–954.

Choi, D., and Kim, J. 2004. Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. *CyberPsychology & behavior* 11–24.

Conole, G., and Culver, J. 2010. The design of cloudworks: Applying social networking practice to foster the exchange of learning and teaching ideas and designs. *Computers & Education* 54:679–692.

DiMicco, J., et al. 2008. Motivations for social networking at work. In *CSCW*. ACM.

Ducheneaut, N., and Moore, R. J. 2004. The social side of gaming: a study of interaction patterns in a massively multiplayer online game. In *CSCW*, 360–369. ACM.

Eysenbach, G. 2008. Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. *Journal of Medical Internet Research* 10.

Gelley, B., and John, A. 2015. Do i need to follow you?: Examining the utility of the pinterest follow mechanism. In *CSCW*, 1751–1762. ACM.

Heiberger, G., and Harper, R. 2008. Have you facebooked astin lately? using technology to increase student involvement. *New Directions for Student Services* 2008(124):19–35.

Hendricks, D. 2014. Are interest-based networks the way of the future? http://www.forbes.com/sites/drewhendricks/2014/10/16/are-interest-based-networks-the-way-of-the-future/.

Jamison, J. 2012. Beyond facebook: The rise of interest-based social networks.

Java, A.; Song, X.; Finin, T.; and Tseng, B. 2007. Why we twitter: understanding microblogging usage and communities. In *WebKDD (SNA-KDD)*, 56–65. ACM.

Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; and Darrell, T. 2014. Caffe: Convolutional architecture for fast feature embedding. In *MM*.

Lee, T. Y., et al. Experiments on motivational feedback for crowdsourced workers. In *ICWSM*.

Linder, R.; Snodgrass, C.; and Kerne, A. 2014. Everyday ideation: All of my ideas are on pinterest. In *CHI*, 2411–2420. ACM.

Macskassy, S. A. 2012. On the study of social interactions in twitter. In *ICWSM*. AAAI.

Musial, K., and Sastry, N. 2012. Social media: are they underpinned by social or interest-based interactions? In *Proceedings of the Fourth Annual Workshop on Simplifying Complex Networks for Practitioners*, 1–6. ACM.

Swamynathan, G.; Wilson, C.; Boe, B.; Almeroth, K.; and Zhao, B. Y. 2008. Do social networks improve e-commerce?: a study on social marketplaces. In *WSON*, 1–6. ACM.

Zhong, C.; Shah, S.; Sundaravadivelan, K.; and Sastry, N. 2013. Sharing the loves: Understanding the how and why of online content curation. In *ICWSM*. AAAI.

Zhong, C.; Salehi, M.; Shah, S.; Cobzarenco, M.; Sastry, N.; and Cha, M. 2014. Social bootstrapping: how pinterest and last. fm social communities benefit by borrowing links from facebook. In *WWW*.

Zhong, C.; Karamshuk, D.; and Sastry, N. 2015. Predicting pinterest: Automating a distributed human computation. In *WWW*.